AD_____

Award Number: W81XWH-04-1-0130

TITLE: Relating Androgen Receptor Conformation to Function in Prostate Cancer Cells

PRINCIPAL INVESTIGATOR: Paul Webb, Ph.D.

CONTRACTING ORGANIZATION: The University of California
San Francisco, California 94143-0962

REPORT DATE: January 2005

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	January 2005	Final (1 Jan 04 - 31 Dec 04)

4. TITLE AND SUBTITLE
Relating Androgen Receptor Conformation to Function in Prostate Cancer Cells

5. FUNDING NUMBERS
W81XWH-04-1-0130

6. AUTHOR(S)
Paul Webb, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of California
San Francisco, California 94143-0962

E-Mail: webbp@itsa.ucsf.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Original contains color plates. All DTIC reproductions will be in black and white.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Androgen Receptor (AR) recruits SRC coactivators in response to stimulation by hormone and other signals, thereby promoting prostate cancer growth and survival. Activated AR adopts two conformations, a head to tail dimer where a coactivator binding site (AF-2) in the C-terminal ligand binding domain (LBD) recognizes a short peptide (FQNLF) in the N-terminal domain of a partner AR and an alternate dimer with AF-2 available for contacts with SRC LxxLL motifs. Here, we set out to identify AR mutants committed to particular conformations and create stable cell lines that express these mutants. As part of an unexpected collaboration, we obtained crystals of AR LBDs in complex with FxxLF and LxxLL peptides and evidence that AR interactions with SRC2 require AF-2 contacts with LxxLL motifs (appendix). We learned how to obtain AR mutants unable to form the head to tail dimer and created vectors suitable for stable AR expression. We created prostate cancer cells that express a tet –repressor protein and are now selecting cells that express AR under tet control and HeLa cells that express human AR or AR mutants. These cells will be used for studies of effects of AR conformation on coactivator recruitment and prostate cancer cell behavior.

14. SUBJECT TERMS
Androgen Receptor, Prostate Cancer Cell Line, Transcriptional Activation, Tet-On, Inhibitor RNA

15. NUMBER OF PAGES
47

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

Androgen Receptor (AR) recruits SRC coactivators in response to stimulation by hormone and other signals, thereby promoting prostate cancer growth and survival. Activated AR adopts two conformations (1). In one, a coactivator binding site (Activation function 2, AF-2) in the C-terminal ligand binding domain (LBD) binds a short hydrophobic peptide (FQNLF) in the N-terminal domain (NTD) of a partner AR, thereby promoting formation of an unusual head to tail dimer. Here, AR must use activation domains within its NTD (AF-1) to bind SRCs. Alternatively, AR can dimerize through contacts between DNA binding domains, with AF-2 available for SRC recruitment. It is not clear whether one, or both, AR conformations drive prostate cancer growth. It is important to find out, strategies that block SRC recruitment could inhibit growth of primary and secondary hormone resistant prostate cancer. **Our goal was to obtain mutant ARs that are committed to one conformation, and to introduce these mutant ARs into cell lines.** Ultimately, we plan to create prostate cancer cells expressing these mutant ARs to understand how AR conformation influences androgen response at the whole cell level.

BODY

We made unexpected progress on Androgen Receptor coactivator preferences. This led to publication on Androgen Receptor Structure (Appended (2)), which will help us identify AR mutants committed to particular active conformations. We experienced initial problems in cell line development, but have made progress towards resolving these issues and in development of alternate strategies.

Task 1. Development of AR and Mutant AR Expression Vectors.
The goal was to create AR mutants committed to particular conformations in a tetracycline-inducible mammalian cell expression vector. We chose the T-rex system from InVitrogen. This is a so-called "Tet-on system" that relies upon coexpression of a tetracycline repressor (from the TR vector) along with the protein of interest (AR, in our case) under control of a vector promoter that contains four binding sites for the Tet-repressor protein (TO vector). AR should be expressed in cells in response to addition of Tetracycline into the medium.

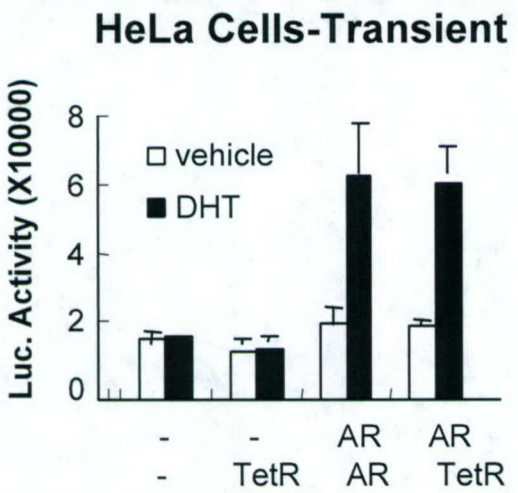


Fig. 1 Transient Transfection with TO-AR Vector. Reporter contained three consensus AREs.

We cloned full length wild type AR cDNA into the TO vector. We confirmed that this vector produced functional AR in transient transfections into HeLa cells, as assessed by production of AR protein observed in Western blot of transfected cell extracts and by assay of AR activity at a transiently cotransfected reporter (Fig. 1). Unexpectedly, coexpression of the Tet-R did not influence AR activity, suggesting that transient expression of high amounts of To-AR vector overcomes the Tet-repressor.

Our choice of AR mutants that are committed to particular conformations was more difficult than we had anticipated. We originally proposed use an AR bearing a mutation within the FQNLF peptide in the NTD. While this AR should be blocked in head to tail dimer formation, and should be committed to the second conformation in which AR AF-2 is available for coactivator binding, recent reports revealed that the FQNLF peptide also recruits cyclin D1, an AR corepressor ((3). We therefore set out to determine whether point mutations in the AR AF-2 surface would distinguish between AR interactions with the FQNLF motif in the NTD, and LxxLL motifs.

Our studies are presented in the appended publication ((2) and are results of a collaboration between three groups, my own, Kip Guy, a pharmaceutical chemist, and Robert Fletterick, an X-ray crystallographer. Briefly, we obtained X-ray crystal structures of the AR LBD in complex with FxxLF peptides and LxxLL peptides derived from SRCs. This enabled us to obtain clear insights into the way that the coactivator binding

surface recognizes both motifs. Contrary to existing models, which suggest that residue K720 is needed for binding to FQNLF but not LxxLL motifs in SRC coactivators (He and Wilson, Mol. Genet. Metab. 2002 75 p293), we show that this residue is absolutely required for binding to both motifs. Instead, and contrary to models that suggest that residue E897 is needed for interactions with both motifs, we find that mutations in this residue can block AR interactions with the FQNLF motif, but not the LxxLL motif. The data also strongly support the idea that AR interactions with SRC2/GRIP1, an important mediator of AR effects on prostate cancer cell growth, require AR AF-2 interactions with SRC2/GRIP1 LxxLL motifs.

Based on these studies, we created several AR mutants in the TO-expression vector:

1) E897 mutations (E897Q, A). These AR mutants should not be able to form the head to tail dimer, but will bind SRCs.

2) AR FQNLF>AA. This mutant will not form the head to tail dimer, but will bind SRCs. We must also consider the possibility that this mutant will not bind efficiently to cyclin D1.

3) ARV716R. This AR mutant will not bind efficiently to FQNLF or LXXLL motifs in SRCs. This should give us a measure of AR AF-1 activity.

4) AR DBD-LBD. This AR lacks AF-1.

We are also searching for AR mutants that are committed to the head to tail interaction. We are replacing the FQNLF motif with a Corepressor Nuclear receptor Interacting motif from N-CoR (CoRnR box). It is known that the AR LBD binds CoRnR boxes when bound to RU486 (4). Our mutant should form the N-C interaction in response to RU486, which usually prevents both the intramolecular interaction and coactivator binding.

Finally, we have created several expression vectors for AR mutants that display extreme superactivity at complex native promoters such as the PSA and probasin genes, but not at the mouse mammary tumor virus promoter. These include an AR with a mutation in the hinge domain that links the LBD to DBD (K632A,K633A) and an AR that lacks SUMOlation sites in the NTD (K385E, K517E). While not part of the original goals of this work we noted that these mutant ARs display a similar promoter-specific phenotype to ARs lacking an intact FQNLF motif, and reasoned that they may also show differential activities as a result of altered AR conformation.

Task 2. Development of LNCaP Cells to Express Human AR mutants.

We experienced unexpected difficulties in Task 2. We transfected the TET-repressor vector into LNCaP cell suing electroporation, and selected for stable transfectants using Blasticidin, which is the resistance marker on the vector. LNCAP cells grew slowly in our hands. It took over six months to isolate Blasticidin-resistant colonies. We found that none of these colonies expressed Tet-repressor mRNA in these cells by standard PCR.

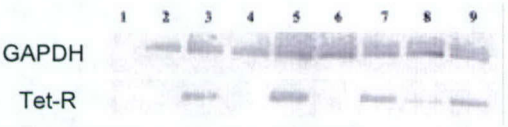


Fig. 2 PCR analysis of Tet R expression in LNCaP cells. 1= Blank. 2=Untransfected LNCaP. 3-9=LNCaP clones.

We repeated the selection using a different transfection technique, Calcium Phosphate instead of electroporation. Once again, the cells grew slowly, and, once again, most colonies lacked detectable tet-repressor mRNA. This time, however, we did detected several LNCAP cell derivatives that express detectable Tet-R mRNA (Fig. 2). We selected one that expressed highest levels for further characterization. We performed killing curves with Zeocin, the selectable marker for the To-AR vector. We also confirmed that we could detect AR activity in this cell line by transient expression. We have stably transfected these cells with expression vectors for TO-AR or several AR mutants, as listed above. Colonies are growing, albeit slowly.

Since we were experiencing problems in obtaining LNCaP cells, we looked for other cell types that would be useful to our aims. It had been shown that AR recruits similar coactivators in prostate cancer cells and human HeLa cells (5). Since HeLa cell derivatives that express the Tet-R are commercially available, we selected AR expressing derivatives of these cells. We have now obtained HeLa cells that stably express most of our key AR mutants, as described in task 1.

Task 3 AR-RNAi.

AR-RNAi has now been successfully used by several groups (see for example, (6)). We have created similar vectors.

Task 4. Characterization of AR expressing Cells.

We have demonstrated that HeLa cells express AR and AR mutants in response to tetracycline induction by western blot of cell extracts (Fig. 3). We can use these cells to determine whether AR conformation affects coactivator association. We expect that our forthcoming LNCaP cells will be used for studies of relationships between AR and cell behavior in forthcoming proposals.

Fig. 3 Tet-inducible expression of AR in stable cells. Note that AR protein is only expressed in transfected cells in response to tet.

KEY RESEARCH ACCOMPLISHMENTS

1) Description of AR AF-2 surface at the atomic level permits construction of AR mutants that are conformation specific.
2) Construction of LNCaP Cells that express the Tet-Repressor.
3) Construction of Stable Human Cell lines expressing AR or AR mutants committed to particular conformations.

REPORTABLE OUTCOMES

The molecular mechanisms of coactivator utilization in ligand dependent transactivation by the androgen receptor. Estebanez-Perpina E, Moore JM, Mar E, Rodrigues ED, Nguyen P, Baxter JD, Buehrer BM, Webb P, Fletterick RJ, Guy RK. J Biol Chem. 2004 Nov 24; [Epub ahead of print]

CONCLUSIONS.

We created AR mutants that distinguish between conformations and have begun to create cell lines that will teach us whether different conformers have different functions.

References

1. **He, B., and Wilson, E.M. The NH(2)-terminal and carboxyl-terminal interaction in the human androgen receptor.** ***Mol Genet Metab*** **75: 293-8 (2002).**
2. **Estebanez-Perpina, E., Moore, J.M., Mar, E., Rodrigues, E.D., Nguyen, P., Baxter, J.D., Buehrer, B.M., Webb, P., Fletterick, R.J., and Guy, R.K. The molecular mechanisms of coactivator utilization in ligand dependent transactivation by the androgen receptor.** ***J Biol Chem*** **(2004).**
3. **Burd, C.J., Petre-Draviam, C.E., Moghadam, H., Wilson, E.M., and Knudsen, K.E. Cyclin D1 binding to the androgen receptor NH2-terminal domain inhibits AF2 association and reveals dual roles for AR co-repression.** ***Mol Endocrinol*** **(2004).**
4. **Hodgson, M.C., Astapova, I., Cheng, S., Lee, L.J., Verhoeven, M.C., Choi, E., Balk, S.P., and Hollenberg, A.N. The androgen receptor recruits nuclear receptor corepressor (N-CoR) in the presence of mifepristone via its amino and carboxy terminii revealing a novel molecular mechanism for androgen receptor antagonists.** ***J Biol Chem*** **(2004).**
5. **Wang, Q., and Fondell, J.D. Generation of a mammalian cell line stably expressing a tetracycline-regulated epitope-tagged human androgen receptor: implications for steroid hormone receptor research.** ***Anal Biochem*** **289: 217-30 (2001).**
6. **Kang, H.Y., Cho, C.L., Huang, K.L., Wang, J.C., Hu, Y.C., Lin, H.K., Chang, C., and Huang, K.E. Nongenomic androgen activation of phosphatidylinositol 3-kinase/Akt signaling pathway in MC3T3-E1 osteoblasts.** ***J Bone Miner Res*** **19: 1181-90 (2004).**

JBC Papers in Press. Published on November 24, 2004 as Manuscript M407046200

Title: The Molecular Mechanisms of Coactivator Utilization in Ligand Dependent Transactivation by the Androgen Receptor

Eva Estébanez-Perpiñá[1], Jamie M. R. Moore[2], Ellena Mar[1], Edson Delgado Rodrigues[3], Phuong Nguyen[3], John D. Baxter[3], Benjamin M. Buehrer[4, 5], Paul Webb[3], Robert J. Fletterick[1], and R. Kiplin Guy[2*]

[1]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143 USA

[2]Departments of Pharmaceutical Chemistry and Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94143 USA

[3]Metabolic Research Unit and Diabetes Center, University of California, San Francisco, CA 94143, USA

[4]Karo Bio USA, Durham, NC 27703 USA

[5]Present Address, Karo Bio AB Huddinge

[*]Corresponding author, email: rguy@cgl.ucsf.edu, phone: 415-502-7051, fax: 415-514-0689.

Running Title: Coactivators in AR Transactivation

Summary

Androgens drive sex differentiation, bone and muscle development, and promote growth of hormone dependent cancers by binding the nuclear androgen receptor (AR), which recruits coactivators to responsive genes. Most nuclear receptors recruit steroid receptor coactivators (SRCs) to their ligand binding domain (LBD) using a leucine rich motif (LxxLL). AR is believed to recruit unique coactivators to its LBD using an aromatic rich motif (FxxLF) while recruiting SRCs to its amino terminal domain (NTD) through an alternate mechanism. Here, we report that the AR LBD interacts with both FxxLF motifs and a subset of LxxLL motifs, and that contacts with these LxxLL motifs are both necessary and sufficient for SRC mediated AR regulation of transcription. Crystal structures of the activated AR in complex with both recruitment motifs reveal that side chains unique to the AR LBD rearrange to bind either the bulky FxxLF motifs or the more compact LxxLL motifs, and that AR utilizes subsidiary contacts with LxxLL flanking sequences to discriminate between LxxLL motifs.

Introduction

The cellular effects of the hormone 5-α-dihydrotestosterone (DHT) are mediated by the androgen receptor (AR), a member of the nuclear hormone receptor superfamily.(1) AR is absolutely required for normal male development, plays a variety of important roles in metabolism and homeostasis in adult men and women, (2,3) and is required for prostate cancer growth. Consequently, AR is a major target for pharmaceutical development and the recognized target for existing prostate cancer therapies, including androgen withdrawal and antiandrogens.(1,4-6) It is nonetheless desirable to obtain new antiandrogens that spare patients from harmful side effects and inhibit AR action in secondary hormone resistant prostate cancer, where AR action becomes sensitized to low levels of androgens or existing antiandrogens.(6,7) Improved understanding of AR signaling pathways will facilitate development of these compounds.

Like most nuclear receptors (NRs), AR activity depends on interactions with members of steroid receptor coactivator (SRC) family.(1,8,9) Several lines of evidence indicate that AR contacts with SRCs are important in prostate cancer. First, androgens promote SRC recruitment to the androgen-regulated prostate specific antigen promoter and this event is inhibited by the antiandrogen flutamide.(10) Second, exogenous SRC2 (GRIP1/TIF2) promotes the androgen-dependent progression from the G1 to S phase in LNCaP prostate tumor cells, in a manner that requires specific AR contact.(10) Third, SRCs often become expressed at high levels in prostate cancers.(5) Finally, AR contacts with SRCs mediate hormone-independent AR signaling in conditions that resemble

secondary prostate cancer.(11,12) Thus, strategies to inhibit AR contacts with SRCs could be useful strategies in blocking prostate cancer cell growth.

For many NRs, overall transcriptional activity stems mostly from the hormone-dependent activation function (AF-2) within the NRs ligand binding domain (LBD), and involves interaction between a conserved hydrophobic cleft on the surface of the LBD and short leucine rich hydrophobic motifs (NR boxes, consensus LxxLL motif) reiterated within each SRC.(13,14) In contrast, current models of AR action suggest that AR activity stems from a potent hormone independent activation function, AF-1, within the AR's N-terminal domain (NTD) and emphasize the role of contacts between NTD and glutamine rich sequences within the SRC C-terminus in SRC recruitment.(15-19) The AR LBD is proposed to bind LxxLL motifs weakly, and instead bind preferentially to aromatic rich motifs that are found within the AR NTD (FQNLF and WHTLF) and AR specific coactivators such as ARA70.(16,20-23) The intramolecular interactions between the LBD and the NTD FQNLF motif promote formation of head to tail dimers (N-C interaction), which render the AF-2 surface unavailable for direct cofactor contacts.(21) Together, the notion that AR AF-2 binds coactivators weakly, and the fact that it will be occluded by the N-C interaction, has led to the suggestion that AR AF-2 does not play an active role in SRC recruitment.

Nonetheless, several lines of evidence suggest that AR AF-2 can contribute directly to coactivator recruitment in some contexts. First, the N-C interaction is required for optimal AR activity at some promoters, including those of probasin, prostate specific antigen, and C3, but not at others, including those of the sex

limiting protein and the mouse mammary tumor virus long terminal repeats (MMTV-LTR).(16) Thus, AF-2 may be available for coactivator contacts in some circumstances. Second, mutation of AR AF-2 recognition sequences within target coactivators inhibits AR coactivation.(16,19,20) Thus, mutation of FxxLF motifs within AR-specific coactivators such as ARA70 blocks their ability to interact with AR and potentiate AF-2 activity. More surprisingly, given the prevailing notion that AR AF-2 contacts with LxxLL motifs are weak, mutation of all three SRC LxxLL motifs inhibits AR coactivation when SRCs are overexpressed, when AR NTD FQNLF and WHTLF motifs are mutated, or when AR acts at promoters such as the MMTV-LTR.

It is important to understand the overall significance of particular AR to coregulator contacts, and the mechanism of these interactions, in order to develop strategies to inhibit AR activity in prostate cancer. In this study, we examine AR AF-2 interactions with target coactivators. Our studies confirm that AR AF-2 binds FxxLF motifs, but also show that AR AF-2 binds a subset of SRC LxxLL motifs with higher affinity and, further, that the same LxxLL motifs are required to mediate AR AF-2 activity. Crystal structures of AR LBD in complex with native FxxLF and LxxLL peptides reveal the structural basis for these unusual coactivator binding preferences, and may suggest new approaches to drug design.

Experimental Procedures

Protein Expression and Purification

AR LBD (residues 663-919) was expressed in *Escherichia coli* (*E. coli)* and purified to homogeneity using a modified version of previously published protocols.(24) Bacterial cell preparations were grown at ambient or lower temperatures to high OD at 600nm (>1.00) in 2XLB supplemented with DHT. AR-LBD protein was expressed by induction with IPTG for 14-16 hours at 15°C before harvest and cell lysis by freeze-thawing and mild sonication. Purification involved an initial affinity chromatography step using a glutathione Sepharose column, followed by thrombin cleavage of the GST affinity tag. Finally cation exchange chromatography with Sepharose SP afforded the purified protein. Our procedures differ from published work in that we use Sepharose SP for the second purification step instead of Fractogel SO_3, which does not retain AR in our experiments.

Peptide Library Synthesis

Coregulator peptides consisting of 20 amino acids with the general motif of **C**XXXXXXX**L**XX**L/AL/A**XXXXXXX were constructed, where C is cysteine, L is leucine, A is alanine, and X is any amino acid. The sequences of all the coregulator peptides were obtained from human isoform candidate genes (SRC1/AAC50305, SRC2/Q15596, SRC3/Q9Y6Q9, ARA70/Q13772,). The peptides were synthesized in parallel using standard fluorenylmethoxycarbonyl (Fmoc) chemistry in 48 well synthesis blocks (FlexChem System, Robbins). Preloaded Wang (Novagen) resin was deprotected with 20% piperidine in

dimethylformamide. The next amino acid was then coupled using 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (2.38 equiv. wt.), Fmoc-protected amino acid (2.5 equiv. wt.), and diisopropylethylamine (5 equiv.wt.) in anhydrous dimethylformamide. Coupling efficiency was monitored by the Kaiser Test. Synthesis then proceeded through a cycle of deprotection and coupling steps until the peptides were completely synthesized. The completed peptides were cleaved from the resin with concomitant side chain deprotection (81% TFA, 5% phenol, 5% thioanisole, 2.5% ethanedithiol, 3% water, 2% dimethylsulphide, 1.5% ammonium iodide) and crude product was dried down using a speedvac (GeneVac). Reversed-phase chromatography followed by mass spectrometry (MALDI-TOF/ESI) was used to purify the peptides. The purified peptides were lyophilized. A thiol reactive fluorophore, 5-iodoacetamidofluorescein (Molecular Probes), was then coupled to the amino terminal cysteine following the manufacturer's protocol. Labeled peptide was isolated using reversed-phase chromatography and mass spectrometry. Peptides were quantified using UV spectroscopy. Purity was assessed using LCMS.

Peptide Binding Assay

Using a BiomekFX in the Center for Advanced Technology (CAT), AR-LBD was serially diluted from 100 μM to 0.002 μM in binding buffer (50 mM Sodium Phosphate, 150 mM NaCl, pH 7.2, 1 mM DTT, 1 mM EDTA, 0.01% NP40, 10% glycerol) containing 150 μM ligand (dihydroxytestosterone) in 96 well plates. Then 10 μL of diluted protein was added to 10 μL of fluorescent coregulator

peptide (20 nM) in 384 well plates yielding final protein concentrations of 50-0.001 μM and 10 nM fluorescent peptide concentration. The samples were allowed to equilibrate for 30 minutes. Binding was then measured using fluorescence polarization (excitation l 485 nm, emission l 530 nm) on an Analyst AD (Molecular Devices). Two independent experiments were assayed for each state in quadruplicate. Data were analyzed using SigmaPlot 8.0 (SPSS, Chicago, Il) and the K_d values were obtained by fitting data to the following equation (y=min+(max-min)/1+(x/K_d)^Hillslope).

GST Pull-Down Assays

Full-length SRC-2 (amino acids 1-1462) and AR NTD-DBD (amino acids 1-660) was expressed in a coupled transcription/translation system (TNT, Promega). AR LBD (amino acids 646-919), or AR LBD mutants, were expressed in *E. coli* strain BL21 as a glutathione S-transferase (GST) fusion protein and attached to glutathione beads according to manufacturer's protocol (Amersham Pharmacia Biotech). Binding assays were performed by mixing glutathione-linked Sepharose beads containing 4 mg of GST fusion protein (estimated by Coomassie Plus protein assay reagent, Pierce) with 2 ml of ^{35} S-labeled SRC-2 or AR NTD-DBD in 20 mM HEPES, 150 mM KCl, 25 mM $MgCl_2$, 10% glycerol, 1 mM dithiothreitol, 0.2 mM phenylmethylsulfonyl fluoride, 20 mg/ml bovine serum albumin and protease inhibitors containing to a final volume of 150 ml. The bead mix was shaken at 4° C for 1.5 h, washed three times in 200 ml of binding buffer. The bound proteins were resuspended in SDS-PAGE loading buffer and were

separated using 10% SDS-polyacrylamide gel electrophoresis and visualized by autoradiography.

Cell Culture and Transfection Assays

HeLa, DU145, and CV-1 cells were maintained in DME H-21 4.5 g/l glucose, containing 10% steroid depleted fetal bovine serum (Gibco), 2 mM glutamine, 50 units/ml penicillin, and 50 mg/ml streptomycin. For transfection, cells were collected and resuspended in Dulbecco's phosphate-buffered saline (0.5 ml/4.5 $x10^7$ cells) containing 0.1% dextrose, and typically 4 µg of luciferase reporter plasmid, 1 µg of AR expression vector or empty vector control, and 2 µg of pCMV-b-galactosidase. Cells were electroporated at 240V and 960 microfarads, transferred to fresh media, and plated into 12-well plates. After incubation for 24 h at 37°C with androgen or vehicle, cells were collected and pellets were lysed by addition of 150 ml of 100 mM Tris-HCl, pH 7.8 containing 0.1% Triton X-100.

For transfections with full length AR, the reporter gene utilized the Mouse Mammary Tumor Virus promoter fused to luciferase. For transfections with GAL-AR LBD, GAL-TR LBD and GAL-CBP fusions, the reporter contained five GAL 4 response elements upstream of a minimal promoter. LUC and b-galactosidase activities were measured using the Luciferase Assay System (Promega) and Galacto-Light Plus beta-Galactosidase Reporter Gene Assay System (Applied Biosystems), according to the manufacturer's instructions.

Crystallization, Structure Determination and Refinement

The complexes of SRC2-2, SRC2-3, SRC3-2 and ARA70 peptides and AR LBD were prepared by mixing at 0°C for 2 hr, with variable ratios of peptide (3 to

10 mM) and protein (at about 4.5 mg/ml). Crystals were obtained by vapor diffusion methods (sitting-drop technique) using crystal screens from Hampton. The protein-peptide complex solution was mixed with the reservoir solution (0.8M Na-Citrate, 0.1 M Tris pH=7.5 or pH=8.0), and concentrated against 300 µl of the reservoir. Crystals appeared after one day and grew to maximal dimensions after four days. After four days these crystals started to crack so new crystallization trials were necessary to find additives that stabilize the crystals. 0.3 µl of either 2.0 M NaCl, 1.0 M LiCl2 or 0.1 M EDTA were added to a 1 µl protein+ -1 µl reservoir drop to stabilize AR LBD crystals at room temperature.

Crystals for either AR-DHT or AR-DHT-peptide were transferred to a new drop containing 10% (v/v) of glycerol for cryoprotection. The crystals were then flash-cooled using liquid nitrogen and measured using the synchrotron radiation at the 8.3.1 beam line at the Advanced Light Source (ALS) (Berkeley). Crystals containing SRC2-3, SRC2-2 and SRC3-2 diffracted to 2.07 Å, 1.66 Å and 2.7 Å, respectively. Co-crystals of ARA70 peptide with AR LBD were also grown and a complete data set was obtained at 2.3 Å resolution. All the crystals belong to space group $P2_12_12_1$ (orthorhombic) and contain one molecule per asymmetric unit.

The diffraction data were integrated and scaled using the computer program ELVES (http://ucxray.berkeley.edu/~jamesh/elves/).(25) Molecular replacement solutions for all AR LBD peptide structures were obtained using rotation and translation functions from Crystallography & NMR Systems (CNS, http://cns.csb.yale.edu/v1.1/).(26)

The first electron maps calculated after the rigid body refinement that followed the molecular replacement displayed clear electron density for the peptides. During the improvement of the protein model, the Fourier maps revealed better electron density for more flanking residues of the peptides. The electron density for the peptide was always modeled as a short α-helix. However, refinement of the SRC2-2 peptide as an α-helix was unsuccessful as such peptide does not adopt such helical conformation on the AR LBD AF2 surface. Further SRC2-2 model building and refinement was not pursued as an α-helix. A composite omit map not including the peptides was calculated in the last steps of refinement for overcoming phase bias for each one of the complexes. This map was calculated omitting 5% of the total model allowing a better tracing of the peptide and also permitted to visualize more residues that were not visible in the 2Fo-Fc map. Model building was done using the program QUANTA (Accelrys Software, http://www.accelrys.com/quanta/) monitored using the free-R factor. Calculation of the electron density maps and crystallographic refinement was performed with CNS using the target parameters of Engh and Huber.(27) Several cycles of model building, conjugate gradient minimization and simulated annealing using CNS resulted in structures with good stereochemistry. A Ramachandran plot shows that most of the residues fall into the most favored or additionally favored regions. The statistics for data collection and refinement of each one of the data sets can be found in table 1.

The structures have been deposited with the Protein Data Bank (PDB) and assigned the following ID numbers: : AR•DHT•SRC2-3; PDB 1T63, RCSB

RCSB022358; AR•DHT•ARA70; PDB 1T5Z, RCSB RCSB022354; AR•DHT•SRC2-2; PDB 1T65, RCSB RCSB022360; AR•DHT•SRC3-2; PDB 1XJ7, RCSB RCSB030414.

Results

AR AF-2 Binds SRC-2 NR Boxes 1 and 3 with High Affinity

To understand the unusual spectrum of AR AF-2 coactivator interactions, we measured binding of the AR-LBD to a library composed of NR boxes from known coactivating proteins including both SRCs and AR specific coactivators (Fig. 1a). Such peptides are known to bind to other NRs with equal affinity to the full-length coactivator.(28) AR-LBD interacted to varying degrees with all of the peptides containing an LxxLL motif tested except the first NR box of ARA70. As expected, AR LBD interacted with FxxLF sequences present in ARA70 and the AR NTD (21,29) fairly strongly with measurable dissociation constants of 33 μM ± 3.3 and 38 μM ±3.8, respectively. Surprisingly, AR also recognized a subset of NR boxes from the SRC family.(30) Specifically, peptides of the first (SRC2-1, K_d = 13 μM ±2.1) and third (SRC2-3, K_d = 15 μM ±1.2) NR boxes of SRC-2 (GRIP1/TiF-2/N-CoA-2) bound strongly to AR, followed in affinity by FxxLF motifs. The second NR box of SRC3 (RAC3/p/CIP/p300/CBP-interacting protein) was also recruited to AR (K_d = 39 μM ± 5). The remaining NR boxes from SRC-1, SRC-2, and NTD weakly interacted with AR either non-specifically or with binding affinities above the assay range (> 40 μM). Control experiments with the same sequences in which LxxLL or FxxLF had been converted to LxxAA or FxxAA revealed the binding was dependent upon the intact triad of hydrophobic amino acids (not

shown). This substitution has been shown previously to abolish interactions with NR.(31)

Pulldown experiments confirmed that the AR LBD bound SRC2 strongly, as opposed to the AR NTD or NTD-DBD (Fig. 1b). Further, AR-LBD interactions with SRC2 were inhibited by mutation of SRC2 boxes 2 and 3 (Fig. 1c), or by increasing concentrations of SRC2-3 peptide (Fig 1d). Thus, AR LBD binds FxxLF motifs, but also binds a subset of classic NR box peptides with comparable or higher affinities. Moreover, the preference of AR for individual LxxLL motifs is different from that observed with other NRs, such as the estrogen receptor (ER) and thyroid receptors (TRs), which bind box 2 in each of the three SRCs with high affinity.(28,32-34)

AR -Dependent Transactivation Requires SRC2 Boxes 1 and 3

Next, we examined the ability of SRC2 to coactivate isolated AR AF-2, and requirements for individual LxxLL motifs in this effect. As expected, a fusion protein containing the AR LBD (amino acids 646-919) linked to the yeast GAL4 DNA binding function conferred androgen-dependent transcriptional activity on a GAL4 responsive reporter in several cell types, and simultaneous expression of SRC2 strongly enhanced AR AF-2 activity (Fig. 2A). Overall, AR AF-2 activity was more potent than that of AR AF-1 in HeLa and DU145, particularly in the presence of SRC2, and about 20-30% as potent as that induced by TR and ERα LBDs, which bind a wider range of SRCs (see supplemental material). As expected from prior results, AF-1 dominates signaling in CV-1 cells, the effects of AF-1 and AF-2 are balanced in DU145 cells, and AF-2 dominates in HeLa

cells.(35,36) Thus, our results are consistent with the notion that AR AF-2 is potent(35,36) and contradict the notion that AR AF-2 has little or no intrinsic activity.

Mutation of individual SRC-2 NR boxes to LxxAA reveals a requirement for boxes 1 and 3 to provide full AR AF-2 activity, both in HeLa (Supplemental Figure) and in DU145 cells (Fig.2B). In contrast, NR box 2 of SRC2 is required to mediate TRβ AF-2 in HeLa (Supplemental Figure), consistent with our own determinations of the affinity of SRC2 NR boxes for TRβ, and with previous results(8,28). Moreover, each mutant SRC showed equivalent ability to enhance activity of CBP AD2, which binds the SRCs at a distinct locus and in a manner that is independent of NR boxes (Supplemental Figure).(8) Thus, the NR box mutations that reduce AR transactivation do not affect other elements of SRC2 activity.

NR boxes also played a role in the ability of SRC2 to coactivate full length AR (Fig. 2C). For these experiments, we utilized an MMTV-LTR driven reporter, because the N-C interaction is dispensable for optimal AR activity at this promoter, and HeLa cells, because AR AF-2 activity is relatively strong in this cell type. Here, SRC-2 enhancement of AR signaling was lessened when the NR boxes were mutated.(17-19,37) In particular, mutation of the third NR box (SRC2-3) abrogated SRC-2 action (Fig. 4). Thus, there is exact congruence between the affinity of particular NR boxes for AR and their requirement for transactivation in the context of the isolated AR LBD and full length AR.

X-Ray Structures of AR-LBD in Complex with Coregulator Peptides Reveal the Atomic Basis for AR Selective Binding to SRC2 NR boxes and ARA 70

To determine how AR binds aromatic rich coactivator domains and a particular subset of SRC NR boxes, we obtained crystal structures of the AR-LBD in complex with ARA70-2, SRC2-2, SRC2-3, and SRC3-2. As expected by analogy with other NR AF-2s, SRC2-3, SRC3-2 and ARA70 peptides bind as a short α-helix into the L-shaped hydrophobic cleft normally utilized by coactivators. On the contrary, the low affinity peptide SRC2-2 was seen to bind to AR LBD AF through an energetically non-favorable conformation that could not be modeled as an α-helix. Comparison of the structures also reveals features that explain the ability of the AR AF-2 to bind to both LxxLL and FxxLF motifs.

The AR LBD crystal structure in complex with the SRC2-3 peptide KENALLRYLLDKDD (14-mer) has been solved to 2.07 Å resolution. Thirteen residues of this peptide are clearly defined in the electron density, and the interaction buries 1322 $Å^2$ of predominantly hydrophobic surface area from both molecules. Our structure shows that SRC2-3 hydrophobic motif binds in nearly the same manner as previously stated in other NRs with LXXLL p160 coactivator motifs.(32,38-40) The residues located N-terminally from the first Leu residue (residue +1) are termed –1, -2, and so on, whereas the residues C-terminal from Leu+1, are termed +2, +3, etc. The core hydrophobic motif of the peptide (residues +1 to +5) forms a short a-helix that binds in the groove formed by helices 3, 4, 5, and 12. The LBD interacts primarily with the hydrophobic face of the SRC2-3 peptide a-helix formed by the side chains of the three LXXLL motif

leucines (Leu923, Leu926 and Leu927). The side chain of Leu923 is embedded within the groove and forms van der Waals contacts with the side chains of Val716, Met734, and Asn738. The side chain of Leu927 is also isolated within the groove and makes van der Waals contacts with the side chains of Gln733 and Met734. The side chain of the second NR box 3 leucine, (Leu926), makes van der Waals contacts with the side chains of Val716 and Met894. The LBD residues implicated in hydrophobic contacts with the peptide are valines 716, 730, and 901, methionines 734, 894, gluthamines 733 and 738, Ile898, and the non-polar parts of Asp731 and Glu 893 and 897.

The main chain carbonyl groups of residues Leu927, Asp930 and Asp931 from the SRC2-3 peptide also interact with Lys720, which is highly conserved in NRs and comprises the upper part of a charge clamp that stabilizes the α-helical NR box peptide conformation. However, contrary to predictions made on the basis of mutagenic analysis of AR surface residues(30), and comparisons with a glucocorticoid receptor (GR)/SRC2-3 structure(39), the SRC2-3 peptide does not form any hydrogen bonds to the second highly conserved charge clamp residue, Glu897 on Helix 12. Instead, the peptide engages in hydrophobic contacts with Glu897, and the distance to the three unpaired amide NH of the peptide helix is 5Å, so electrostatic stabilization is possible. The peptide also engages in hydrogen bonding to seven water molecules in its vicinity. Residue Asp928 located at position +6 adopts two different conformations. However, neither Asp928 (+6) nor Arg924 (+2) interact with charged residues on the AR surface that comprise a second charge clamp, again contrary to predictions made on the

basis of a GR/SRC2-3 structure.(39) Nonetheless, the SRC2-3 peptide displays clear electron density in the current structure for five residues N-terminal to the core hydrophobic motif and for four more residues C-terminal to the same motif, therefore displaying significantly greater electron density than any other NR box peptide in complex with a NR LBD to date.

The AR LBD crystal structure in complex with the SRC3-2 peptide HKKLLQLLT (9-mer) has been solved to 2.7 Å resolution. All nine residues of this peptide are clearly defined in the electron density, and the interaction buries 1052 $Å^2$ of predominantly hydrophobic surface area from both molecules. Our structure shows that SRC3-2 hydrophobic motif binds in nearly the same manner as previously stated for SRC2-3. The LBD residues implicated in hydrophobic contacts with the peptide are valines 716, 730, methionines 734, 894, Ile898, and the non-polar parts of Glu 897 and Lys720, unexpectedly. SRC3-2 peptide is shorter C-terminally than SRC2-3, and does not make any hydrogen bonds with Lys720. Surprisingly, another basic residue, Arg726 adopts in this complex the C-terminal capping role stabilizing the peptide a-helix. This polar interaction is not present in the other peptide- AR LBD complexes described in this paper. This crystal structure shows traceable electron density for 6 residues located at the protein N-terminus that correspond to some residues of the hinge region of AR, this is the first time that such residues are visible in an electron density. Those residues are in a random coiled-coil conformation.

The AR LBD complex with the SRC2-2 peptide comprises the following sequence, KHKILHRLLQDSS (13-mer). Despite the fact that the crystal of SRC2-

2 diffracted to 1.66 Å, the electron density that accounts for the peptide was more difficult to interpret and discontinuous suggesting that its affinity for AR-LBD is weak. It was surprising to state that SRC2-2 adopts two different conformations. The first was very similar to the SRC2-3 peptide and was modeled as a short α-helix. However a second conformation, more similar to a coiled coil could be interpreted and refined (referred as the non-canonical conformation). In the SRC2-3 like conformation, interpretable electron density starts at the first leucine of the SRC2-2 peptide and finishes at Gln928. Building this peptide from the Box3 conformation leaves only correctly placed within the weak electron density Leu923, His924 bulges out of the density and only the main chain returns to the electron density for the following Arg925, Leu926, Leu927, and Gln 928. On the other hand, if NR box 2 is built and refined as a random coil, interpretable and continuous electron density starts at residue His920 until Leu926. From all these residues, only Leu923 is completely defined, the rest of 6 residues only the main chain is defined in the electron density, leaving the side chains unseen. NR box 2 in Box3-like conformation buries 850 $Å^2$, whereas NR box 2 in random coil conformation buries 792 $Å^2$ of predominantly hydrophobic surface area from both molecules.

In the Box 3-like conformation, the side chain of Leu923 is embedded within the groove and forms van der Waals contacts with the side chains of Leu712, Asn738, Met894 and Ile898. The side chain of Leu927 makes van der Waals contacts with the side chain of Met734. The side chain of the second NR box 2 leucine, (Leu926), makes van der Waals contacts with the side chain of Val716.

The LBD residues implicated in hydrophobic contacts with the peptide are Val716, methionines 734 and 894, Gln738, Ile898, and the non-polar part of Glu893. The residue Leu926 interacts with **Lys720**, through its main chain carbonyl group. In the non-canonical conformation, Leu926 also interacts with **Lys720**, through its main chain carbonyl group. NR box 2 peptide does not form any hydrogen bonds to the second highly conserved charge clamp residues, **Glu897**, in either conformation. However, His920 could be bonded to Glu893. Except for three N-terminal residues that are disordered, the position and interactions of the ARA70 FxxLF peptide with the AR surface more closely recapitulate the binding mode observed in structures of ternary complexes of SRC LxxLL motifs with hormone bound NR LBDs (Fig. 3a and 3c).(32,38-40) The triad of aromatic side chains (FxxLF) that forms the hydrophobic face of the coactivator helix fits tightly into a deep narrow pocket. In addition, charged residues at either end of the cleft, Glu897 and Lys720, cap the helix (the "charge clamp"). The fully engaged interaction is manifested in the tight binding of this coactivator and its strong transactivation.

The AR LBD Charge Clamp Plays Coregulator Selective Roles in Transactivation and Binding.

One unexpected feature of our crystal structures is that the two residues that comprise the canonical AR LBD charge clamp (Lys720 on helix 3 and Glu 897 on helix 12) interact differently with FxxLF and LxxLL peptide backbones. While previous studies suggested that Glu897 was absolutely required for SRC binding, our structures revealed that Glu897 is fully engaged with the carbamyl backbone

of the FxxLF peptide, but not that of the LxxLL peptide. Similar arrangements were also observed in crystals of the AR LBD in complex with artificial FxxLF and LxxLL peptides derived from phage display.(41)

To understand the apparent discrepancy between the reported requirement for Glu897 in AR activity and its lack of contact with the LxxLL motif of SRC2-3 in the crystal structure, we examined the effects of a series of charge clamp mutations on isolated AR AF-2 activity *in vivo* (Fig. 4a) and coregulator binding *in vitro* (Fig. 1a). As expected, a mutation within the upper charge clamp residue (Lys720Ala) inhibited AR AF-2 activity (Fig. 4a) and prevented the recruitment of SRC2(Fig. 4c). The reversal of the normal negative charge at Glu 897 by introduction of a positive charge (Glu897Lys, Glu897Arg) had the same effect, probably due to repulsion of the charged NR box.(30,35) However, AR-LBDs bearing mutations that neutralized or lessened electrostatic potential at Glu897 (Glu897Ala, Glu897Gln) retained significant AF-2 activity, especially in the presence of SRC2 (Fig 4a).(9) These same mutants had no discernable effects upon recruitment of SRC2 (Fig 4c), and a modest effect on recruitment of the AR NTD. This is in keeping with the effects of the Glu897Gln mutation on NR box peptide recruitment (Fig 1a). Western blotting of cell extracts confirmed that these differences in transcriptional activity were not related to differential expression of the AR LBD mutants. Thus, the lower charge clamp residue (Glu897) is dispensable for SRC-2 binding but required for ARA70 binding, exactly paralleling the requirement for this residue observed in both of our crystal structures.

Discussion

In this paper, we examined the binding of AR AF-2 to a range of target motifs within potential AR coactivators, confirmed the functional consequence of these interactions and determined how AR AF-2 binds selectively to particular motifs. Our results confirm that AR AF-2 recognizes FxxLF motifs derived from the AR NTD and ARA70 with moderate affinity (<40uM), but also show that AR binds some LxxLL motifs, particularly SRC2-1 and SRC2-3, with higher affinity (<10μM). The discovery that AR AF-2 binds strongly to selected LxxLL motifs is surprising, but several lines of evidence confirm the importance of these interactions. Thus, bacterially expressed AR-LBD binds SRC2 strongly, as compared to TRβ AF-2 and AR AF-1, and these interactions are dependent upon NR boxes. Moreover, isolated AR AF-2 activates transcription relatively strongly, and does so in a manner that is potentiated by SRC2 and dependent upon SRC2-1 and SRC2-3. Finally, SRC2 LxxLL motifs were required for coactivation of full length AR; at least at the MMTV promoter. Thus, AR AF-2 binds FxxLF motifs, but can also make important contacts with a subset of coregulator LxxLL motifs. AR therefore has the potential to activate transcription in an analogous manner to other NRs.

To understand the unusual selectivity of AR AF-2 for target coactivator motifs, we solved the structures of the AR-LBD in complex with an FxxLF motif derived from ARA70 and both high affinity (SRC2-3) and low affinity (SRC2-2) AR interacting motifs. Our structures indicate that the ARA70 FxxLF motif occupies a similar position to those of other coregulator NR box peptides in

complex with LBDs of other NRs. Comparisons of each of the ternary complexes with each other, and with our own structures of AR in the absence of an associated peptide (not shown), reveal a striking rearrangement of the AF-2 surface that explain the ability of AR to accommodate the bulky hydrophobic side chains of the FxxLF motifs. Movements of Lys720, Met734, and Glu897 create the deeper pockets and enhanced electrostatics allowing the binding of the ARA70 peptide (See Fig. 3). Similar rearrangements were also observed in crystals of AR LBD in complex with artificial FxxLF and LxxLL peptides derived from phage display.(41) Of these residues, Met734 is relatively unique amongst the NR superfamily, and only conserved at an equivalent position within the glucocorticoid receptor LBD. Thus, the presence of Met734 probably explains the unique capacity of the AR AF-2 surface to bind accommodate motifs with bulky hydrophobic side chains.

Crystal structures of AR-LBD in complex with SRC2-3 and SRC2-3 suggest an alternate explanation for the ability of AR AF-2 to discriminate between different LxxLL motifs. The SRC2-3 and SRC2-2 LxxLL motifs, by contrast to the ARA70 FxxLF motif and a variety of NR box peptides in complex with a variety of NR LBDs, are translated by about 2Å in the cleft, towards helix 3. Overall, this unusual positioning disrupts the electrostatic stabilization characteristic of most NR/NR box interactions, likely explaining reduced AR binding to most LxxLL motifs. However, for SRC2-3, the high degree of negative charge in the four residues following the motif (sequence DKDD) interacts with positively charged patches on the receptor surface. In fact, these portions of the structure are better

ordered than in all previous NR-coactivator complexes, and are not visible in AR LBD structures with the SRC2-2 peptide, which binds the AR-LBD with lower affinity. This influence offsets suboptimal electrostatics and explains the selective binding of AR AF-2 to SRC2-3. Thus, AR discriminates between cofactor NR box motifs by making auxiliary contacts outside of the core LxxLL motif. Interestingly, the ARA70 peptide is also relatively well ordered, about 12 of 15 amino acids are visible in our crystal structure. While it has been previously suggested that NR LBDs may discriminate between target motifs by contacting residues that flank the hydrophobic LxxLL core,(28,31) our studies provide the first description of a structural basis for this effect.

AR AF-2 has the potential to participate in transcriptional activation in several ways, but the relative importance of different modes of AR AF-2 action are not yet clear. The N-C interaction is required for optimal AR action at a variety of androgen-regulated promoters, including those of prostate specific genes such as PSA and probasin, suggesting that AF-2 mediates intramolecular interactions in these contexts. We predict that AR AF-2 could participate in coactivator binding in several contexts, including in the presence AR specific coactivators that contain FxxLF motifs, in conditions of SRC2 overexpression, and at promoters that resemble the MMTV-LTR. The requirement for AR AF-2 in growth of prostate cancer cells has not been rigorously addressed, but it is interesting to note that SRC2 enhancement of the androgen-dependent G1 to S transition in LNCaP prostate tumor cells is dependent upon the integrity of the SRC2 NR box region (which binds AF-2) and independent of the SRC2 C-terminus (which binds

AR AF-1).(10) Perhaps AR AF-2 contacts with SRC LxxLL motifs will prove to be relevant for cell cycle progression.

In conclusion, AR has a potent AF-2 that drives the cell's expression program by binding FxxLF motifs and selected LxxLL motifs. The receptor uses the same general coactivator binding mechanisms as other NRs, by providing a dimorphic cleft that facilitates interaction with aromatic amino acids in addition to leucines. The ability of the AR surface to rearrange to interact with FxxLF motifs is unique among transcription factors and represents a gain of function relative to other structurally defined interactions in the family. Most NRs are unable to accommodate bulky sidechains in the binding domains of the coactivators, and the dyadic recognition of AR has enabled development of more complex control mechanisms involving the NTD and the use of specialized subsets of coactivators. Most importantly, the new function *does not* come at the cost of a loss of ability to interact productively with SRCs. AR AF-2 interactions with SRCs are likely to be physiologically relevant, particularly in certain forms of prostate cancer.

References

1. Lee, H. J., and Chang, C. (2003) *Cell Mol Life Sci* **60**, 1613-1622
2. Liu, P. Y., Death, A. K., and Handelsman, D. J. (2003) *Endocr Rev* **24**, 313-340
3. Legros, J. J., Charlier, C., Bouillon, G., and Plomteux, G. (2003) *Ann Endocrinol (Paris)* **64**, 136
4. Gregory, C. W., He, B., Johnson, R. T., Ford, O. H., Mohler, J. L., French, F. S., and Wilson, E. M. (2001) *Cancer Res* **61**, 4315-4319
5. Culig, Z., Klocker, H., Bartsch, G., and Hobisch, A. (2002) *Endocr Relat Cancer* **9**, 155-170
6. Santos, A. F., Huang, H., and Tindall, D. J. (2004) *Steroids* **69**, 79-85
7. Balk, S. P. (2002) *Urology* **60**, 132-138; discussion 138-139
8. Ding, X. F., Anderson, C. M., Ma, H., Hong, H., Uht, R. M., Kushner, P. J., and Stallcup, M. R. (1998) *Mol Endocrinol* **12**, 302-313
9. Berrevoets, C. A., Doesburg, P., Steketee, K., Trapman, J., and Brinkmann, A. O. (1998) *Mol Endocrinol* **12**, 1172-1183
10. Shang, Y., Myers, M., and Brown, M. (2002) *Mol Cell* **9**, 601-610
11. Gregory, C. W., Fei, X., Ponguta, L. A., He, B., Bill, H. M., French, F. S., and Wilson, E. M. (2004) *J Biol Chem* **279**, 7119-7130
12. Blaszczyk, N., Masri, B. A., Mawji, N. R., Ueda, T., McAlinden, G., Duncan, C. P., Bruchovsky, N., Schweikert, H. U., Schnabel, D., Jones, E. C., and Sadar, M. D. (2004) *Clin Cancer Res* **10**, 1860-1869
13. Needham, M., Raines, S., McPheat, J., Stacey, C., Ellston, J., Hoare, S., and Parker, M. (2000) *J Steroid Biochem Mol Biol* **72**, 35-46
14. Leo, C., and Chen, J. D. (2000) *Gene* **245**, 1-11
15. Alen, P., Claessens, F., Schoenmakers, E., Swinnen, J. V., Verhoeven, G., Rombauts, W., and Peeters, B. (1999) *Mol Endocrinol* **13**, 117-128
16. He, B., Lee, L. W., Minges, J. T., and Wilson, E. M. (2002) *J Biol Chem* **277**, 25631-25639
17. Ma, H., Hong, H., Huang, S. M., Irvine, R. A., Webb, P., Kushner, P. J., Coetzee, G. A., and Stallcup, M. R. (1999) *Mol Cell Biol* **19**, 6164-6173
18. Christiaens, V., Bevan, C. L., Callewaert, L., Haelens, A., Verrijdt, G., Rombauts, W., and Claessens, F. (2002) *J Biol Chem* **277**, 49230-49237
19. Powell, S. M., Christiaens, V., Voulgaraki, D., Waxman, J., Claessens, F., and Bevan, C. L. (2004) *Endocr Relat Cancer* **11**, 117-130
20. He, B., Minges, J. T., Lee, L. W., and Wilson, E. M. (2002) *J Biol Chem* **277**, 10226-10235
21. He, B., and Wilson, E. M. (2002) *Mol Genet Metab* **75**, 293-298
22. Zhou, Z. X., He, B., Hall, S. H., Wilson, E. M., and French, F. S. (2002) *Mol Endocrinol* **16**, 287-300
23. He, B., Kemppainen, J. A., and Wilson, E. M. (2000) *J Biol Chem* **275**, 22986-22994
24. Matias, P. M., Donner, P., Coelho, R., Thomaz, M., Peixoto, C., Macedo, S., Otto, N., Joschko, S., Scholz, P., Wegg, A., Basler, S., Schafer, M., Egner, U., and Carrondo, M. A. (2000) *J Biol Chem* **275**, 26164-26171

25. Holton, J., and Alber, T. (2004) *Proc Natl Acad Sci U S A* **101**, 1537-1542
26. Brunger, A. T., Adams, P. D., and Rice, L. M. (1998) *Curr Opin Struct Biol* **8**, 606-611
27. Engh, R. A., and Huber, R. (1991) *Acta Crystallographica* **A47**, 392-400
28. Darimont, B. D., Wagner, R. L., Apriletti, J. W., Stallcup, M. R., Kushner, P. J., Baxter, J. D., Fletterick, R. J., and Yamamoto, K. R. (1998) *Genes Dev* **12**, 3343-3356
29. Bourguet, W., Andry, V., Iltis, C., Klaholz, B., Potier, N., Van Dorsselaer, A., Chambon, P., Gronemeyer, H., and Moras, D. (2000) *Protein Expr Purif* **19**, 284-288
30. He, B., and Wilson, E. M. (2003) *Mol Cell Biol* **23**, 2135-2150
31. McInerney, E. M., Rose, D. W., Flynn, S. E., Westin, S., Mullen, T. M., Krones, A., Inostroza, J., Torchia, J., Nolte, R. T., Assa-Munt, N., Milburn, M. V., Glass, C. K., and Rosenfeld, M. G. (1998) *Genes Dev* **12**, 3357-3368
32. Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard, D. A., and Greene, G. L. (1998) *Cell* **95**, 927-937
33. Ribeiro, R. C., Apriletti, J. W., West, B. L., Wagner, R. L., Fletterick, R. J., Schaufele, F., and Baxter, J. D. (1995) *Ann N Y Acad Sci* **758**, 366-389
34. Moore, J. M., Galicia, S. J., McReynolds, A. C., Nguyen, N. H., Scanlan, T. S., and Guy, R. K. (2004) *J Biol Chem*
35. Slagsvold, T., Kraus, I., Bentzen, T., Palvimo, J., and Saatcioglu, F. (2000) *Mol Endocrinol* **14**, 1603-1617
36. Wang, Q., Lu, J., and Yong, E. L. (2001) *J Biol Chem* **276**, 7493-7499
37. Bevan, C. L., Hoare, S., Claessens, F., Heery, D. M., and Parker, M. G. (1999) *Mol Cell Biol* **19**, 8383-8392
38. Shiau, A. K., Barstad, D., Radek, J. T., Meyers, M. J., Nettles, K. W., Katzenellenbogen, B. S., Katzenellenbogen, J. A., Agard, D. A., and Greene, G. L. (2002) *Nat Struct Biol* **9**, 359-364
39. Bledsoe, R. K., Montana, V. G., Stanley, T. B., Delves, C. J., Apolito, C. J., McKee, D. D., Consler, T. G., Parks, D. J., Stewart, E. L., Willson, T. M., Lambert, M. H., Moore, J. T., Pearce, K. H., and Xu, H. E. (2002) *Cell* **110**, 93-105
40. Pike, J. W., Yamamoto, H., and Shevde, N. K. (2002) *Adv Ren Replace Ther* **9**, 168-174
41. Hur, E., Pfaff, S. J., Payne, E. S., Gron, H., Buehrer, B. M., and Fletterick, R. J. (2004) *PLoS Biol* **2**, E274
42. DeLano, W. L. (2002), DeLano Scientific, San Carlos, CA, USA

Figure Legends

Figure 1. The androgen receptor ligand binding domain (AR LBD) binds a subset of steroid receptor coactivator (SRC) nuclear receptor interaction motifs (NR Boxes). Panel A: Sequences of relevant NR Boxes and relative equilibrium affinities of these NR boxes for binding to AR LBD and a mutant AR LBD (E897Q) in which one charge clamp residue has been neutralized. The binding affinities were determined using fluorescence polarization with fluorescently labeled NR box peptides. The coregulator peptides are listed in the left column where SRC1-1, SRC1-2, and SRC1-3, represent the first, second, and third NR box in SRC1, respectively. Each color represents a unique K_d range as defined by the legend in the bottom right-hand corner. For coregulator peptides that displayed saturated binding curves with AR, the actual K_d values are listed. The gray color represents conditions where some interaction of coregulator peptides with AR was observed, however saturating binding curves were not achieved in the protein concentration range studied. Panel B: Pulldown of SRC2 by GST fusions of AR domains. Panel C: Effects of mutation of NR boxes of SRC2 on the Pulldown by the GST fusion of the AR LBD. SRC2 (2,3m) indicates the SRC2 protein where NR boxes 2 and 3 have been mutated from LxxLL to LxxAA. Panel D: Competition for binding of SRC2 by NR box peptides during a Pulldown of SRC2 by the GST fusion of the AR LBD.

Figure 2. Transcriptional activation by AR, AR-NTD, and AR-LBD constructs and the enhancement of activation by SRC constructs. Panel A. Transcriptional

activation of a GAL4-luciferase reporter construct by fusions of GAL4 DNA binding domain with AR-NTD or LBD domains in three cell lines. In all cell lines, AR-LBD induces signaling in response to DHT and this effect is enhanced by expression of SRC2. The level of AR-NTD driven expression varies from cell line to cell line but remains constant in the presence or absence of both DHT and SRC2. Panel B. The effects of mutation of SRC2 NR boxes 1 through 3 upon signaling by GAL4-AR LBD constructs from a GAL driven luciferase reporter. Mutation of SRC2-1 and SRC2-3 both significantly reduce potentiation of transactivation by AR. These mutational effects correlate with the observed relative affinities of the NR boxes for the receptor. Panel C. Activation of transcription at an MMTV-luciferase reporter by full length AR and the effects of co-expression of SRC2 and mutants. Mutation of SRC2-3 significantly reduces potentiation of transactivation by AR.

Figure 3. (A-H) Associations of the AR-LBD with coactivator domains determined by X-ray crystallography. Close-up views of the interaction between ARA70, SRC2-3, SRC2-2, and SRC3-2 peptides with AR LBD AF2. The nuclear receptor AF-2 transactivation function is ascribed to a surface exposed hydrophobic cleft comprised of residues from helices 3 (H3, dark blue), 5 (H5, pale blue) and 12 (H12, red), as can be clearly seen in the bottom figures (E-H). (A-H) The helix backbone of peptides from ARA70 (RETSEKFKLLFQSYN) (left red), SRC2-3 (KENALLRYLLDKDD) (middle left, yellow), and SRC3-2 (HKKLLQLLT) (middle right, orange) are shown, and the non-helical SRC2-2

peptide backbone (KHKILHRLLQDSS) (right, green) can be seen. AR LBD is represented by a solid semi-transparent surface (grey) on the top figures (A-D). The side chains of the motif hydrophobic residues Phe+1/Leu+1, Leu+4 and Phe+5/Leu+5 of the peptides are shown as stick models. Helix 12, with its Glu897 side chain stabilizes the N-terminus of the ARA70 peptide, but not those of the SRC peptides. On H3, the side chain of K720 is shown capping the C-terminus of ARA70 and SRC2-3 peptides (E and F). B. The side chains of the AR LBD residues contacting the peptides are depicted as stick models. ARA70: The triad compressed by the Phe aromatic side chains and Leu+4(FxxLF) fits tightly into a deep narrow pocket comprised of Val716 and Val730, Met734, Ile737, and the hydrophobic segment of Glu893. The Leu side chains of SRC2-3 and SRC3-2 fit loosely into a flat hydrophobic pocket comprised of the side chains of three valines, 716, 730, and 901, methionines 734, 894, glutamines 733 and 738, and Asp731 and Glu897. The accommodation of the bulkier Phe residues of ARA70 is accompanied by the rearrangements of Met734, Glu897, and Lys720 predominantly (indicated by grey dots on the AR's surface representation). (D and H) SRC2-2 does not bind to AR LBD AF2 in an helical conformation and apart from Leu+1, the rest of the peptide cannot be superimposed to the other SRC peptides shown in this paper. All the figures were generated with Pymol (42).

Figure 4. Role of the binding pocket and charge clamp residues of the AR LBD AF-2 in interaction with cofactors and potentiation of transcriptional activation by a GAL4-AR LBD construct. Panel A. Removing the charge at K720 or reversing the charge at E897 (the positive and negative ends of the "charge clamp" that stabilizes helix dipole for the NR box) markedly reduces the potentiation of transcriptional activation by GAL AR LBD by SRC2 in HeLa cells. However, neutralization of the charge at E897 has modest effects on transcriptional activation. Panel B. Western blot demonstrating that all E897 mutants are expressed at similar levels in HeLa cells during the transactivation experiments. Panel C. As expected from the peptide binding data (Figure 1a), neutralization of charge at E897 has no discernable effect upon the interaction of SRC2 as measured by GST Pulldown. Similarly, there is a modest reduction in binding of AR NTD by E897Q. However, reversal of charge (E897K) strongly reduces binding of both SRC2 and AR NTD.

TABLE 1. Statistics for Data Collection and Refinement

	AR-SRC2-3	AR-SRC2-2 (non-helical)	AR-SRC3 (RAC3)	AR-ARA70
Molecules/ASU	1	1	1	1
Space group	P212121	P212121	P212121	P212121
Cell constants a/b/c (Å)	54.49/67.37/ 70.52	55.60/67.58/ 69.32	53.06/66.83/ 71.07	55.68/66.42/ 68.25
Resolution (Å)	2.07	1.66	2.7	2.3
Reflections Measured	393765	511617	375686	458173
Unique reflections	16416	35221	17753	13713
Overall Completeness (%)	97.2	91.7	90	92.8
Outermost shell Completeness (%)	94.3	88.0	83.8	85.2
R merge (%)[a]	4.4	6	5.5	5
Reflections used refinement	15915	32260	6151	10881
Resolution range (Å)	24-2.07	25-1.66	25-2.7	24-2.3
R factor (%)[b]	19.8	21.1	25.3	22.8
R free (%)[c]	23.2	24.8	31.5	25.8
Number of water molecules	160	361	100	106
Matthews Coefficient	2.157	2.116	2.100	2.104
Solvent content (%)	43	42	41.5	40
Ramachandran plot most favored (%)	93	92	82	92
Ramachandran plot allowed (%)	7	7	17	8

ASU=asymmetric unit

[a] R merge (%)= $\sum_{hkl} |<I>-I| / \sum_{hkl} |I|$

[b] R factor (%)= $\sum_{hkl} ||Fo| - |Fc| / \sum_{hkl} |Fo|$

[c] R free set contained 5% of total data

Acknowledgements

We thank Eugene Hur, Luke M. Rice, Elena Sablin, and Maia Vinogradova for useful discussions; James Holton, Corie Ralston, and ALS beamline staff for assistance in data collection and processing. The Prostate Cancer Foundation, NIH (DK58080, RJF and RKG; DK51281, JDB; SPORE NCI CA8952, EEP), and DOD (DAMD17-01-1-0188, JRM and PC030607, PW) supported this work.

Competing Interests and Materials Requests

JDB has proprietary interests in, and serves as a consultant and Deputy Director to, Karo Bio AB, which has commercial interests in this area of research. The other authors declare they have no competing interests. Requests for materials should be addressed to the corresponding author.

FIGURE 1

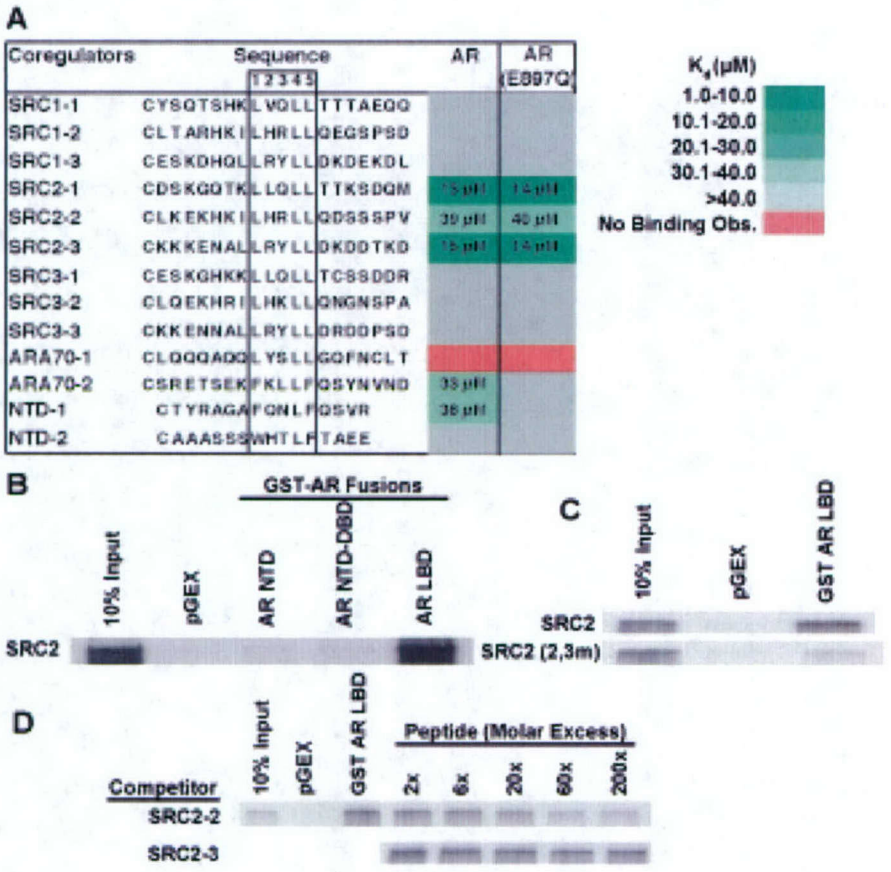

A
Coregulators
Sequence
1 2 3 4 5
AR
AR (E897Q)
SRC1-1 CYSQTSHKLVQLLTTTAEQQ
SRC1-2 CLTARHKILHRLLQEGSPSD
SRC1-3 CESKDHQLLRYLLDKDEKDL
SRC2-1 CDSKGQTKLLQLLTTKSDQM
SRC2-2 CLKEKHKILHRLLQDSSSPV
SRC2-3 CKKKENALLRYLLDKDDTKD
SRC3-1 CESKGHKKLLQLLTCSSDDR
SRC3-2 CLQEKHRILHKLLQNGNSPA
SRC3-3 CKKENNALLRYLLDRDDPSD
ARA70-1 CLQQQADQLYSLLGQFNCLT
ARA70-2 CSRETSEKFKLLFQSYNVND
NTD-1 CTYRAGAFQNLFQSVR
NTD-2 CAAASSSWHTLFTAEE
K_a (μM)
1.0-10.0
10.1-20.0
20.1-30.0
30.1-40.0
>40.0
No Binding Obs.
B
GST-AR Fusions
10% Input
pGEX
AR NTD
AR NTD-DBD
AR LBD
SRC2
C
10% Input
pGEX
GST AR LBD
SRC2
SRC2 (2,3m)
D
10% Input
pGEX
GST AR LBD
Peptide (Molar Excess)
Competitor
2x
6x
20x
60x
200x
SRC2-2
SRC2-3

FIGURE 2

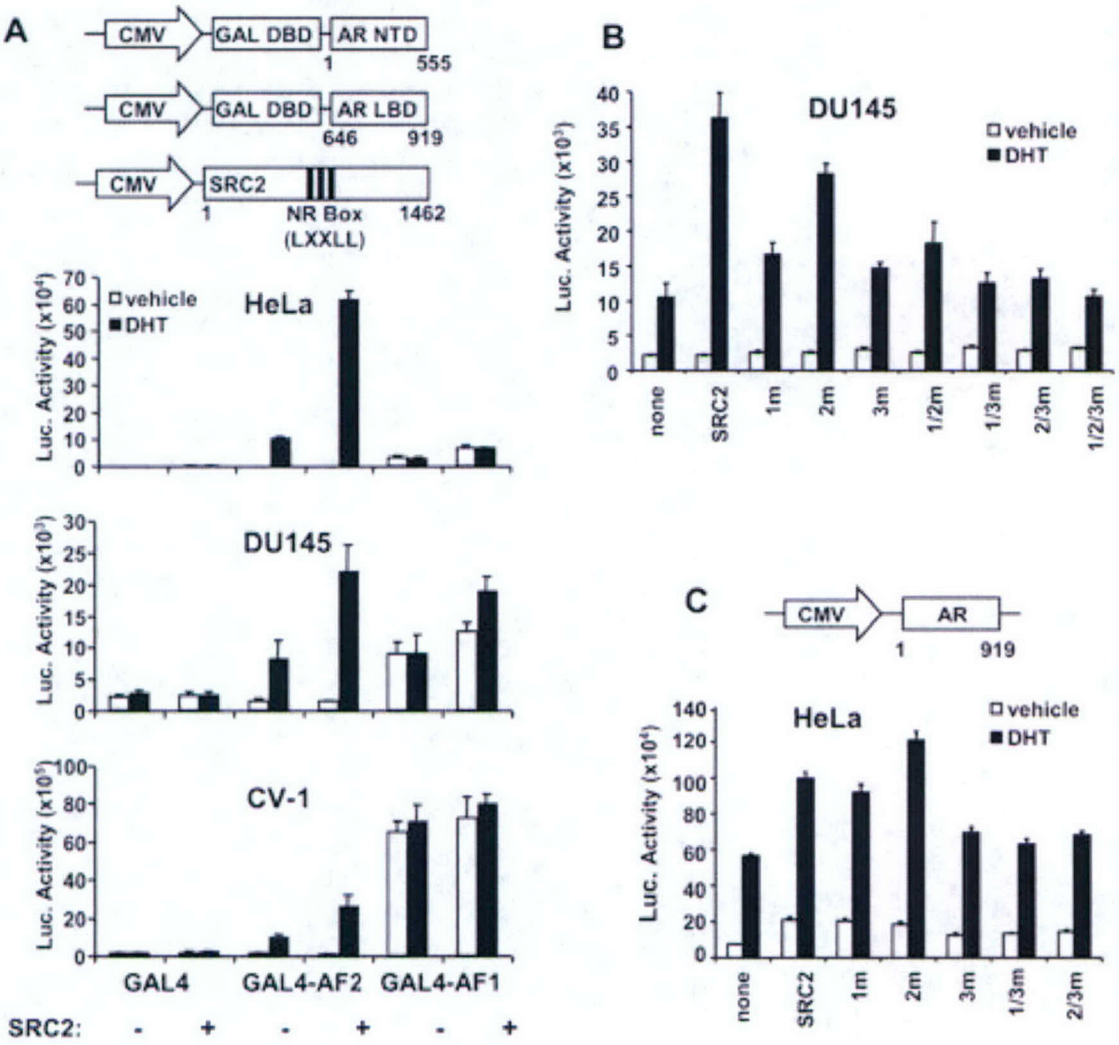

A
CMV
GAL DBD
AR NTD
1
555
CMV
GAL DBD
AR LBD
646
919
CMV
SRC2
1
NR Box
(LXXLL)
1462
vehicle
DHT
HeLa
Luc. Activity (x10⁴)
DU145
Luc. Activity (x10³)
CV-1
Luc. Activity (x10⁵)
GAL4
GAL4-AF2
GAL4-AF1
SRC2: - + - + - +
B
DU145
vehicle
DHT
Luc. Activity (x10³)
none
SRC2
1m
2m
3m
1/2m
1/3m
2/3m
1/2/3m
C
CMV
AR
1
919
HeLa
vehicle
DHT
Luc. Activity (x10⁴)
none
SRC2
1m
2m
3m
1/3m
2/3m

FIGURE 3

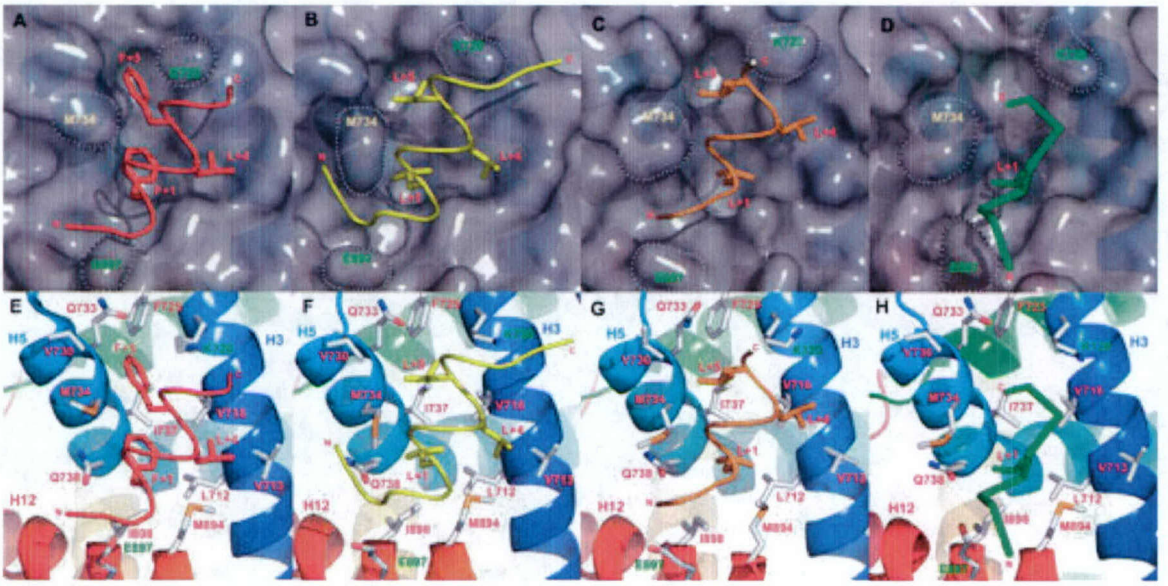
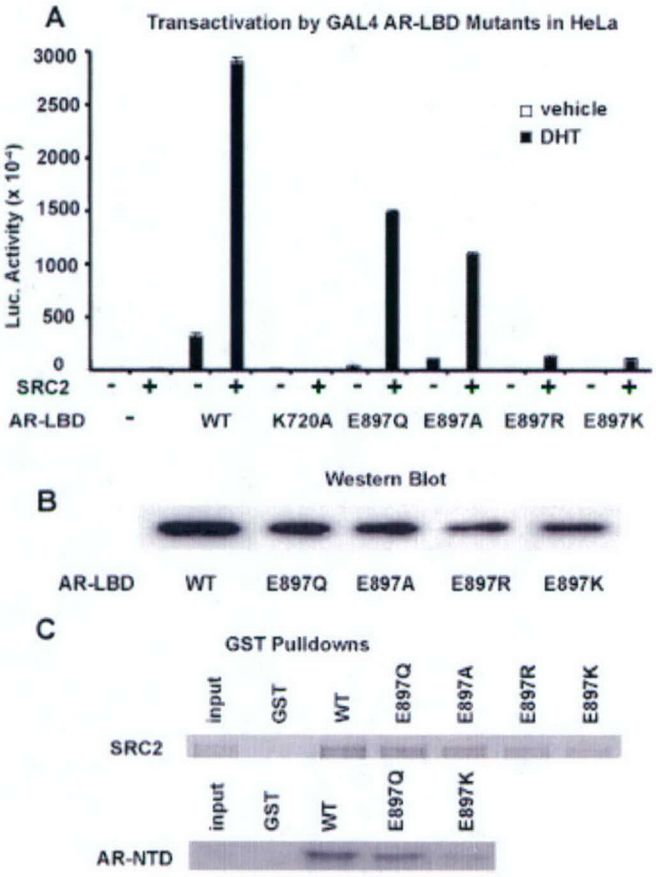
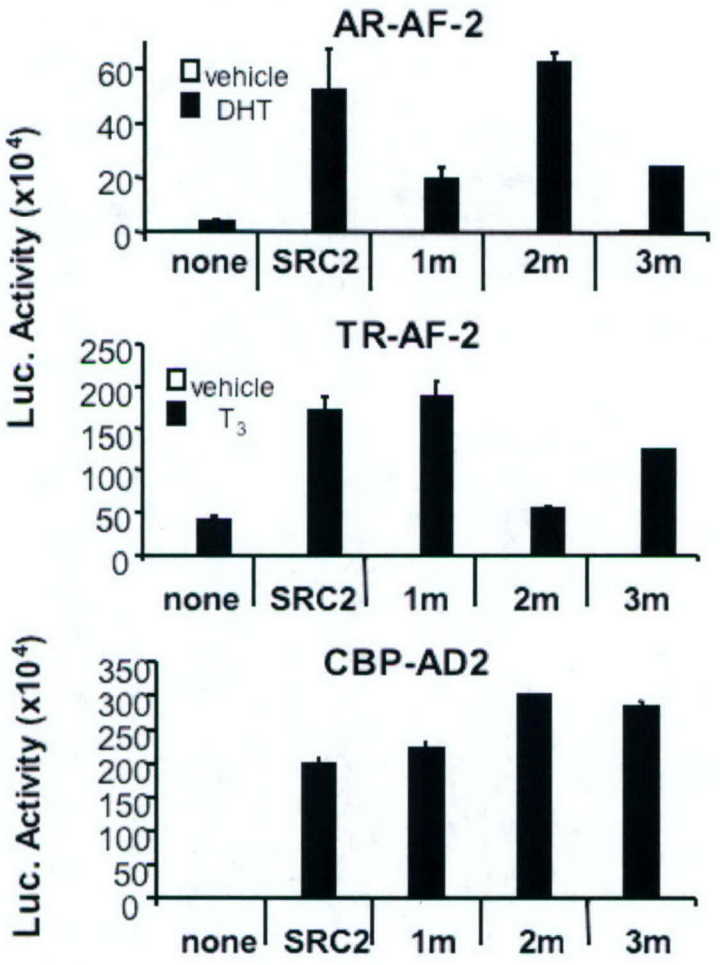


FIGURE 4


A
Transactivation by GAL4 AR-LBD Mutants in HeLa
Luc. Activity (x 10⁻³)
3000
2500
2000
1500
1000
500
0
□ vehicle
■ DHT
SRC2 - + - + - + - + - + - + - +
AR-LBD - WT K720A E897Q E897A E897R E897K
B
Western Blot
AR-LBD WT E897Q E897A E897R E897K
C
GST Pulldowns
input GST WT E897Q E897A E897R E897K
SRC2
input GST WT E897Q E897K
AR-NTD

Supplemental Figure



Supplemental Figure. The effects of mutation of SRC2 NR boxes 1 through 2 upon signaling by GAL4-AR and GAL4-TR LBD constructs from a GAL driven luciferase reporter. Mutation of SRC2-1 and SRC2-3 both significantly reduce potentiation of transactivation by AR while not affecting transactivation by TR. Contra wise, Mutation of SRC2-2 significantly reduces transactivation by TR but not by AR. These mutational effects correlate with the observed relative affinities of the NR boxes for the respective receptors. As expected, none of the

mutations significantly reduces transactivation by CBP, which interacts with a distinct locus on the SRC2 molecule, C-terminal to the LxxLL motifs.